

機車交通事故之醫療費用預測模型研究

蔡興國¹ 廖彥琪² 黃維信³ 朱薇薇⁴

摘要

在台灣，事故傷害位居十大死因中的第五位，其中因「交通運輸」或「機動車交通事故」死亡者，佔全體事故傷害中的五成。而在 85-88 年間交通事故申報健保費用一年最高可達 6300-9900 萬。本研究應用邏輯斯迴歸分析(Logistic Regression Analysis)建構機車交通事故所造成的高醫療費用之機率模型，評估哪些因素對於造成高醫療費用具有顯著之影響。研究資料來自中央健保局資料庫及交通部道路交通事故資料庫，因為駕駛人發生交通事故時的各項相關資料（如時間、道路種類、路面情況、…等）記載於道路交通事故資料庫，而其就醫所造成的健保給付則必須至健保資料庫才能得知。本研究針對民國 90 年間機車交通事故受傷之駕駛人，且事故傷害後第一次住院者進行分析，總計 21,687 筆案件，將醫療費用分為二元應變項，其值為高或低二類。由於在許多實際社會、個體選擇及醫學診斷的問題上，二元的結果相較於連續數值的結果來的具有意義。統計分析結果發現下列情況下所發生之交通事故比其他情況來的多：日間、晴天、路面乾燥、上午 8~11 點、市區道路、直線道路、18~24 歲、有適當駕照、男性、重型機車駕駛、有戴安全帽、交叉路口、未保持行車安全距離、機車與機車、交通號誌不正常及正常速限情形。從所建立的預測模型分析，機車交通事故中，對於醫療費用支出的高低有顯著的影響因素為：發生時間、事故類型(機車與機車、人與機車等)、駕駛資格、有無穿戴安全帽、性別、年齡、肇事因素(如不當行駛、超速失控等)。

壹、前言

隨著工業時代的來臨，民眾生活型態的改變，傳染病已被事故傷害與慢性疾病所取代而成為造成民眾死亡的主因之一。近三十年來，事故傷害已經成為公共衛生界的重要議題之一，西元 1949 年有專家首先提出事故傷害與傳統傳染病一樣，有爆發流行、季節變化、長期趨勢、人時地聚集等流行病學特徵，因此，也

本研究為交通部運輸研究所 MOTC-IOT-94-SDB004 計畫之部份成果

¹ 大葉大學資訊工程系副教授。

² 中臺醫護技術學院醫護管理研究所研究生。

³ 中臺醫護技術學院資訊管理系助理教授。

⁴ 中臺醫護技術學院資訊管理系講師。

可以使用相同的方法來研究[1]。根據世界衛生組織(World Health Organization)2000年的統計，全球約有490萬人死於「事故傷害」，故世界衛生組織於2000年3月成立「事故傷害及暴力防制部門(Department for Injuries and Violence Prevention)」，目的是要領導全球進行事故傷害的調查、提供監控系統之協助、界定工作重點和評估指標，以及促進部門間的合作，並且訂定2004年之世界健康日(World Health Day)主題為「道路安全(Safe road)」。

在台灣，事故傷害位居十大死因中的第五位，其中因「交通運輸」或「機動車交通事故」死亡者，佔全體事故傷害中的五成[2]。根據[3]研究指出85~88年間交通事故申報健保費用一年最高可達6300~9900萬。交通事故傷害同樣也是台灣地區年輕人死亡的主要原因，根據衛生署統計，民國91年台灣地區年輕人(15~24歲)因事故傷害而死亡的總人數為1,170人，死亡率佔該年齡層死亡率之52.14%。[4]曾針對事故傷害死亡的變化與國外作過比較，結果發現台灣地區的年輕人，是以機動車所引發的交通事故傷害最為嚴重，需要優先處理。另外也有學者研究指出機動車所引發的死亡率、潛在生命年數損失及貨幣價值較腦血管疾病為高，故其建議對於預防機動車交通事故傷害方面，應多分配資源以減少社會的損失[5]。

在國外，希臘在1980~1993年間因交通意外事故死亡的人數每年增加8.8%，而其所造成的醫療費用也呈逐年增加的趨勢[6]。[7]曾統計韓國在1995年交通意外事故的次數256,052件，其中死亡的數字是12,653超出了先前的記錄相當於22.6%；在美國1990年[8]，因交通事故直接或間接損失，每年約1,480億元，這顯示出各國之間都因交通事故的影響使得醫療費用相對性的成長。

而在一般統計的定量研究中，迴歸分析是最常使用的統計方法，然而在許多情況下，迴歸分析會受到限制。例如，應變項可分為兩類，連續資料(continuous data)或離散資料(discrete data)，當應變項屬於類別資料(categorical variable)而不是連續資料(continuous variable)時，線性迴歸分析就不適用。反之，當應變項為連續資料分析時，可以使用一般線性迴歸模型(General Linear Regression Model)，能自一群自變項中預測應變項後，廣泛地應用於各個領域的研究中。在實際生活中有許多社會科學的問題，所產生的資料大都是偏向類別型態而非連續型態，因此，適用於邏輯斯迴歸(Logistic Regression)分析。因為邏輯斯迴歸是針對二元應變項(0或1)而產生出其模型，運用邏輯斯迴歸的統計分析方法可以處理大量資料。將大量資料依其變項分類後，可以有效處理分析其機率，而各個變項所得到的模型，經由不同的加權指標合併後，可以得到一個很好的預測模型。

邏輯斯迴歸(Logistic Regression)在統計分析應用方面已經有很多年了，但是從1967年後才逐漸普遍。使用邏輯斯迴歸分析最主要目的是為了要找出類別型態的應變項，與一連串的自變項間的關係，因此和迴歸分析中最大的差別在於應變數型態的不同[9]。由於在許多實際社會、個體選擇及醫學診斷的問題上，二元的結果相較於連續數值的結果來的具有意義。本研究使用健保資料庫及道路交通

事故資料庫，透過邏輯斯迴歸分析建構機車交通事故所造成的高醫療費用之機率模型，透過模型評估何種因素具有造成高醫療費用之負擔，提供警察機關及其他相關單位使用，做為預防的工具。

在一般或商業管理方面，[10]利用 1975~1983 年間 38 家的破產企業和正常企業為樣本，先檢定自變數是否符合常態分配的假設，再決定採用區別分析或邏輯斯模型，研究結果發現以邏輯斯模型估計之參數較具一致性，且邏輯斯模型優於區別分析。[11]以分層系統抽樣法並配合電話訪問，取得樣本，利用邏輯斯分析，建立選舉投票傾向推估模式。[12]在台灣上市公司財務評等中羅吉斯模式與判別分析模式之比較中，可解決參數的溢位及偏誤的問題。[13]利用 Logistic Regression 模式針對國內某家金融機構信用卡部門之客戶為研究對象，進行分析持卡人可能發生正常繳款與逾期繳款之特徵因素。[14]使用邏輯斯迴歸，探討影響台灣地區銀行承做擔保放款契約的決定因素。

醫療保健方面，[15]利用邏輯斯迴歸預測兩種醫療技術所選擇的醫療效果。[16]利用 Logistic Regression，分析老年人中西醫併用與只利用西醫之現況及其相關特性分佈與複向就醫之醫療資源使用差異，並探討台灣地區老年人是否有所謂的「逛醫師行為」。[17]以卡方檢定以及多元邏輯斯分析(multiple logistic regression)，探討醫師的科別與職級、藥品的類別與劑型以及環境時空等因素與處方異常的相關性。[18]針對我國全民健康保險的財務結構一直受到各界的高度重視，其中又以醫療費用最受矚目，本研究使用資料探勘技術來建置申報案件核減與否的分類預測模型，並比較邏輯特迴歸與類神經網路所建構的模型。

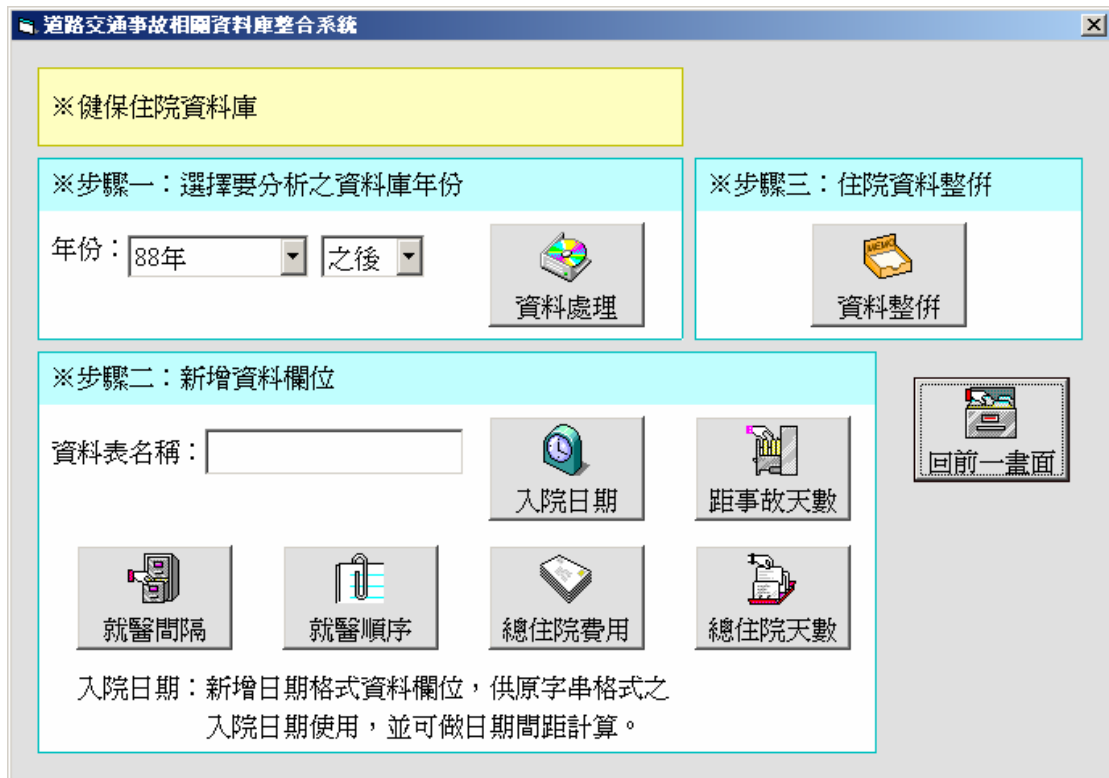
交通事故方面，[19]研究採用民國 90 年台閩地區 A1 類(造成人員當場或 24 小時內死亡)及 A2 類(造成人員受傷或超過 24 小時死亡)交通事故資料，分別建構汽、機車單一車輛事故駕駛人死亡勝算羅吉斯迴歸模式，模式中應變數為代表事故嚴重程度之二元變數(死亡或未死亡)，自變數取自肇事駕駛人的屬性、道路及環境、車輛、單一車輛事故類型等相關因素，除進行模式的適合度檢定外，並經由顯著性統計檢定，確認影響駕駛人高死亡勝算之因子，且以勝算比概念解釋自變數對事故嚴重程度產生之作用。[20]利用 logistic regression 建構出駕駛者特性及行為對於造成交通事故相關因子的模型。[21]使用 logistic regression 研究是否年齡或性別(或二者)是汽車在鄉村道路或是高速公路造成交通事故的影響因素。[22]利用 cross-sectional 問卷調查，針對高雄四所大學的大學生騎機車發生意外事故及機車事故傷害的相關因子做探討，而其相關因子和事故間的關係則是利用對數迴歸分析來探討(logistic regression)。

本研究使用邏輯斯迴歸分析除了讓我們更加清楚自變項與應變項之間的關係外，還可以對不同的屬性做分類，用來對未來發生的事情做預測，並利用資料庫中原有的資料建構出預測模型，以對機車交通事故之醫療費用做預測分析。

貳、材料與方法

一、資料庫整合

資料來自中央健保局資料庫及交通部道路交通事故資料庫(如圖 1、圖 2 所示)，首先，先針對機車事故而受傷之駕駛者第一次住院之資料，透過駕駛者之身份證字號，與道路交通事故資料庫做連結及查詢。而連結合併後所產生之資料檔，選擇可能影響健保給付及部份負擔醫療費用之項目使為自變項(因素)進行統計，並分析自變數之適當性(是否影響醫療費用)及定義高醫療費用標準，研究對象則由此資料檔為分析對象。



資料來源：交通部運輸研究所(2004)

【圖 1 健保資料庫】



資料來源：交通部運輸研究所(2004)

【圖 2 道路交通事故資料庫】

二、研究對象

本研究對象為本國國籍且設籍於臺灣地區之居民，在民國 90 年間機車交通事故受傷之駕駛人，且事故傷害後第一次住院者，總計 21,687 筆案件。

三、研究變項

自變項之選擇係依據相關研究[2,19,24]歸類分析而定，共有發生時間、道路種類、道路型態、年齡、有/無安全帽、肇事因素、事故類型、性別等 8 項。另外，再加入駕駛資格、光線、輕/重型機車、路面狀態、號誌動作、天候、速限等相關變項，以期增加模型預測之準確率，而經由卡方檢定結果，這些變項亦有顯著相關性。

四、統計方法

本研究以醫療費用作為二元應變項，其值為高或低二類。醫療費用的高低是以 21,687 筆案件的總醫療費用取其平均值 38,154 元為分類臨界值，低於(含)此值者歸為低醫療費用，反之則屬高者。

使用 SPSS12.0 做描述性統計，計算各研究變項之次數百分比。其次運用卡方統計量，檢定各自變項與醫療費用間是否具有其相關性。若是 p-value 大於 0.05 則表示自變項與應變項間，無顯著相關性；反之，當 p-value 小於 0.05 則表示自變項與應變項間，有顯著相關性。

在統計資料中，有些是數量資料，有些是質量資料或是類別資料，而在日常生活中，相當容易碰到兩種以上的類別問題，例如消費者對商品的偏好程度，民眾對政府的施政滿意度等。對於這些問題，大致上只能區分出偏好的順序，很難以具體的數據來表示其偏好程度，因此，對於這一類的問題可以使用卡方檢定來分析。

五、預測模型

本研究以logistic regression模型為基本架構，主要使用於應變項為二元資料，其中p介於0~1之間[27]，其方法如下：

$$\begin{aligned}\log\left(\frac{p}{1-p}\right) &= \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \\ &= \beta_0 + \sum_{i=1}^n \beta_i X_i \\ p &= \frac{e^{\beta_0 + \sum \beta_i X_i}}{1 + e^{\beta_0 + \sum \beta_i X_i}}\end{aligned}$$

其中， $\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$ 為線性函數； X_n 為自變項； β_n 為邏輯斯迴歸係數，及n為自變項之數目。

參、結果

一、描述性分析：

表1為各自變項之基本統計分析描述，在發生時間方面，在上午8點及11點所佔的比例較高，凌晨0~6點所佔的比例較低。在道路種類方面，市區道路所佔比例最多，有62.9%，在國道所佔之比例最低，為0%。在道路型態方面，在交叉路與直路部份所佔比例最高，分別為37.6%及36.8%，而在平交道及坡路部份所佔比例最低，分別為0%、0.8%。在年齡方面，18~24歲所佔之比例最高，30.1%，17歲以下則最低，為4.3%。在資格方面，有適當駕照者所佔之比例較高，有71.3%。在光線方面，日間自然光線所佔之比例較高，為62.7%，而在晨或暮光及夜間(或隧道)無照明最低，分別為4.9%與4.1%。在性別方面，男性方面所佔之比例較女生高，為62.3%。在當事者方面，重型機車方面所佔之比例較高，為68.2%。在安全帽方面，有戴安全帽所佔之比例較高，為89.2%。在事故位置方面，交岔路口附近所發生意外事故的比例最高為59.7%，而在交通島及行人穿越道或人行道方面所佔比例最低，分別為0.5%、0.6%。

表1 統計分析

自變項	類別	發生次數	百分比(%)	自變項	類別	發生次數	百分比(%)
發生時間	0~6	2,214	10.2	肇事因素	不當行駛	2,972	13.7
	6~8	2,873	13.2		不當轉彎	1,286	5.9
	8~11	3,494	16.1		超速失控	2,486	11.5
	11~14	3,132	14.4		搶越行人穿越道	1,722	7.9
	14~17	4,006	18.5		未保持安全行車距離	5,587	25.8
	17~20	3,198	14.7		未注意他車(人)安全	128	0.6
	21~23	2,770	12.8		吸食危禁物後駕駛	1,378	6.4

自變項	類別	發生次數	百分比(%)	自變項	類別	發生次數	百分比(%)
道路種類	國道	7	0.0	事故類型	疲勞(患病)駕駛	78	0.4
	省道	1,709	7.9		裝載及攜件	1,189	5.5
	縣道	1,859	8.6		裝載_肇事逃逸	4,715	21.7
	鄉道	1,373	6.3		其他	146	0.7
	市區道路	13,639	62.9		人與汽(機)車	487	2.2
	村里道路	2,812	13.0		車與車	19,408	89.5
	其他	288	1.3		汽(機)車本身	1,792	8.3
道路型態	平交道	4	0.0	路面狀態	冰雪,油滑,泥濘	68	0.3
	三岔路	3,988	18.4		濕潤	2,229	10.3
	交岔路	8,144	37.6		乾燥	19,390	89.4
	單路—其他	799	3.7	速限	0	31	0.1
	彎曲路及附近	603	2.8		4	1	0.0
	坡路	170	0.8		5	1	0.0
	直路	7,979	36.8		10	3	0.0
年齡	12~17歲	933	4.3		15	11	0.1
	18~24歲	6,529	30.1		20	45	0.2
	25~34歲	3,819	17.6		25	77	0.4
	35~44歲	3,129	14.4		30	766	3.5
	45~54歲	2,989	13.8		35	5	0.0
	55~64歲	2,033	9.4		40	10,820	49.9
	65歲以上	2,255	10.4		42	1	0.0
資格	無適當駕照	6,222	28.7		45	9	0.0
	有適當駕照	15,465	71.3		48	1	0.0
光線	日間自然光線	13,595	62.7		50	6,819	31.4
	晨或暮光	1,057	4.9		60	2,552	11.8
	夜間(或隧道)有照明	6,135	28.3		64	1	0.0
	夜間(或隧道)無照明	900	4.1		70	533	2.5
性別	男	13,511	62.3		80	7	0.0
	女	8,176	37.7		81	1	0.0
當事者	重型	14,800	68.2		90	1	0.0
	輕型	6,887	31.8		100	2	0.0
安全帽	有	19,336	89.2				
	無	2,351	10.8				
事故位置	交岔路口附近	12,956	59.7				

自變項	類別	發生次數	百分比(%)	自變項	類別	發生次數	百分比(%)
	交通島	102	0.5				
	快車道	4,691	21.6				
	機車專用道	452	2.1				
	慢車道	1,997	9.2				
	行人穿越道及人行道	121	0.6				
	其他	1,368	6.3				
號誌動作	正常	5,998	27.7				
	不正常	15,689	72.3				
天候	強風, 暴雨	46	0.2				
	霧或煙	19	0.1				
	雨	1,800	8.3				
	陰	2,295	10.6				
	晴	17,527	80.8				

在肇事因素方面，未保持安全行車距離所佔之比例較高，為 25.8%，而在未注意他車(人)安全及疲勞(患病)駕駛所佔之比例最低，分別為 0.6%、0.4%。在事故類型方面，車與車之間所產生的意外事故最高，佔 89.5%，人與汽(機)車所產生的意外事故最低，佔 2.2%。在路面狀態方面，路面乾燥所佔的比例最高，為 89.4%，而在冰雪，油滑或泥濘的路面所佔的比例最低，為 89.4%。在號誌動作方面，號誌不正常運作的情況下所佔比例最高，為 72.7%。在天候方面，晴天的狀況下，所佔比例最高，為 80.8%，在強風、暴雨或是霧、煙的情況下所佔的比例最低，分別為 0.2%、0.1%。在速限方面，正常速限下所佔之比例最高，為 49.9%。

二、卡方 (χ^2) 檢定：

利用卡方分配依據機車交通事故之發生時間、道路種類、道路型態、年齡、駕駛資格、光線、性別、當事者、有/無安全帽、事故位置、肇事因素(如不當行駛、超速失控等)、事故類型(機車與機車、人與機車等)、路面狀態、號誌動作、天候、速限等變項，分別檢定其與醫療費用高低之相關性，檢定結果所有 p 值都小於 0.05，顯示各自變項與預測變項具顯著相關性。

三、Logistic 迴歸：

使用 Logistic Regression 分析建立預測模型，以上述之自變項來預測，而將其中「時間」為 7 類，「光線」分為 4 類，「道路種類」分為 7 類，「道路型態」分為 6 類，「當事者」分為 2 類，「資格」分為 2 類，「年齡」分為 7 類，「性別」分為 2 類，「安全帽」分為 2 類，「事故位置」分為 7 類，「肇事因素」被分為 11 類，「事故類型」分為 3 類，「路面狀態」分為 3 類，「號誌動作」分為 2 類，「天候」分為 5 類，「速限」分為 21 類。

研究結果在模型解釋能力上先假設所求出的 Logistic 迴歸式對預測沒有

幫助，若其顯著機率 p-value 比顯著水準 $\alpha (0.05)$ 小，就表示對所求之結果有幫助；研究結果所求出的顯著機率為 0 小於 $\alpha (0.05)$ ，表示此模型在解釋能力上有顯著幫助。

而在判斷此模型的適合度檢定方面，先假設所求出的 Logistic 迴歸式是適合的，若其顯著機率 p-value 比顯著水準 $\alpha (0.05)$ 大，就表示對所求之結果有幫助；研究結果所求出的顯著機率為 0.443 大於 $\alpha (0.05)$ ，表示此模型在適合度方面有顯著的幫助。

羅吉斯迴歸模式之係數解釋，是以勝算比(Odds Ratio, OR)來表明，亦即 $Exp(\beta)$ 之值。而勝算比是以相對的觀念來描述，即以虛擬變數(Dummy Variable)為參照，加以相對解釋。因此，勝算比(OR)為：

$$OR = \frac{p(x)}{1-p(x)} = Exp(\beta_0 + \beta_1 \chi_1 + \beta_2 \chi_2 + \dots + \beta_n \chi_n)$$

亦即發生機率相對於不發生的機率強度。對勝算比取對數後得：

$$\ln \frac{p(x)}{1-p(x)} = \beta_0 + \beta_1 \chi_1 + \beta_2 \chi_2 + \dots + \beta_n \chi_n$$

若欲比較各類別發生的機率，則需將依變數還原到勝算比(Odds Ratio)，需要進行逆轉換，即是將勝算比取指數(Exponent)，迴歸式中 β 係數則轉為：

$$\frac{p}{1-p} = Exp(\beta_0 + \beta_1 \chi_1 + \beta_2 \chi_2 + \dots + \beta_n \chi_n)$$

下列表 2 為 Logistic Regression 與各變項之整理結果，此預測模型，在分界值(cut-value)為系統內鍵值 0.5 時，其準確預測率為 74.3%。

表 2 Logistic Regression 分析

自變項	速 限 (km/hr)									
類別	4	5	10	15	20	25	30	35	40	42
β	-3.940	-2.986	1.085	-1.589	-0.664	-0.989	-0.611	-1.348	-0.667	-2.889
Exp(B)	0.019	0.050	2.959	0.204	0.515	0.372	0.543	0.260	0.513	0.056
類別	45	48	50	60	64	70	80	81	90	100
β	-1.508	4.860	-0.624	-0.532	-4.326	-0.651	0.010	-3.348	-4.807	-0.986
Exp(B)	0.221	129.021	0.536	0.587	0.013	0.521	1.010	0.035	0.008	0.373
自變項	天 候				性別	當事者	安全帽		資格	
類別	強風， 暴雨	霧 或 煙	雨	陰	男性	重型	無安全帽		無駕駛資格	
β	0.433	0.535	0.025	0.062	0.474	0.006	0.208		0.261	
Exp(B)	1.542	1.707	1.025	1.064	1.606	1.006	1.231		1.298	
自變項	年 齡 (歲)						事故類型		路面狀態	
類別	18~24	25~34	35~44	45~54	55~64	65 以上	機 車 與 車	汽 (機) 車 本 身	濕潤	乾燥
β	0.089	0.094	0.393	0.426	0.635	0.532	0.252	1.287	0.313	0.313

Exp(B)	1.093	1.099	1.481	1.532	1.887	1.703	0.585	1.796	1.367	1.367	
自變項	道 路 種 類							號 誌 動 作			
類別	國道	省道	縣道	鄉道	市區道路	村里道路	不正常				
β	0.518	0.052	-0.121	-0.118	-0.113	-0.098	0.046				
Exp(B)	1.679	1.054	0.886	0.889	0.893	0.906	0.046				
自變項	道 路 型 態										
類別	平交道	交岔路- 三岔路		交岔路	單路-其他		彎曲路及附近		坡路		
β	1.374	0.067		0.081	0.043		0.210		-0.061		
Exp(B)	3.952	1.069		1.085	1.044		1.234		0.941		
自變項	事 故 位 置										
類別	交通島	快車道		機車專用道		慢車道		行人穿越道 及人行道		其他	
β	0.373	0.127		0.046		0.022		0.151		0.148	
Exp(B)	1.452	1.135		1.047		1.022		1.163		1.159	
自變項	肇 事 因 素										
類別	不 當 行 駛	不 當 轉 彎	超 速 失 控	搶 越 行 穿 道	未 持 全 車 離	保 安 行 距	未 注 意 他 車 (人) 安 全	吸 食 危 禁 物 後 駕 駛	疲 勞 (患 病) 駕 駛	裝 載 及 攜 件	裝 載 肇 事 逃 逸
β	0.120	0.016	-0.004	0.210	0.017	0.227	0.202	-0.292	0.139	-0.048	
Exp(B)	0.120	0.016	-0.004	0.210	0.017	0.227	0.202	-0.292	0.139	-0.048	
自變項	時 間						光 線				
類別	0~6	6~8	8~11	11~14	14~17	17~20	日間自然 光線	晨 或 暮 光	夜間(或) 隧道有照 明		
β	0.241	-0.064	-0.014	-0.042	-0.185	-0.099	-0.112	-0.165	-0.125		
Exp(B)	0.241	-0.064	-0.014	-0.042	-0.185	-0.099	0.894	0.848	0.882		

肆、結論與討論

本研究使用道路交通事故資料庫與健保資料庫之整合，以 Logistic Regression 技術建構完成機車交通事故與醫療費用高低之預測模式，使用變項來判別醫療費用之高低。從表 2 的結果中可以發現，以時間中的 21~23 時為參照，其中在凌晨 0~6 時相對於其他類別有較高之影響性。在道路型態中以直路為參照，其中在平交道方面的影響性較高於其他類別。而在年齡的部份是以 12~17 歲為參照，而在 55~64 歲方面的影響性高於其他類別。在事故位置方面以交岔路口附近為參照，其中又以交通島部份的影響性高於其他變項。在肇事因素方面是以其他為參照，肇事因素中各類別間之影響性都較為平均。而在事

故類型中以人與汽(機)車為參照，以汽(機)車本身對於其他類別較具有影響性。在速限中以 0 為參照，而在時速 48km/hr 的影響性遠大於其他類別。

本研究所採用醫療費用高低之標準是依據 90 年度機車交通事故之總醫療費用平均數做為劃分，因此，在各個年度所制定之醫療費用高低會隨著當年度之醫療費用而有所不同，本研究所建構之預測模型可隨各個年度醫療費用所訂定之高低標準而產生不同之準確率，而所選入之變項，對於預測之準確率能達其顯著性。

本研究以 Logistic Regression 所建構的預測模型中所選定的自變項的確會對醫療費用造成較大部份的影響。日後可再以資料探勘的相關技術(如 Decision Tree, Artificial Neural Networks)進行研究分析，以提升預測正確率。

參考文獻

- 1、 呂宗學、李孟智、周明智(1998)。事故傷害研究資料收集與分析的相關問題。中山醫學雜誌，9(1)，25-33。
- 2、 行政院衛生署(2003)。臺灣地區青年主要死亡原因-民國 91 年衛生統計資訊網，見：<http://www.doh.gov.tw/statistic/data/死因摘要/91年/表6.xls>
- 3、 黃勝堅、曹昭懿(2003)。台灣地區外傷事故之健保醫療費用分析：西元 1996~1999 年。台灣醫學，7，861-870。
- 4、 呂宗學、李孟智、周明智(2001)。臺灣年輕人之傷害死亡率：變化型態與國際比較。中山醫學雜誌，12，1-10。
- 5、 陳立慧、林茂榮、王榮德(1993)。機動車交通事故之死亡率、潛在生命年數損失及其貨幣價值。中華衛誌，12(4)，368-379。
- 6、 Skalkidou A., Petridou E., Papadopoulos F. C., Dessypris N., Trichopoulos D., 1999. Factors affecting motorcycle helmet use in the population of Greater athens, Greece. Injury Prevention ,5, 264-267.
- 7、 Sohn S. Y., Shin. H., 2001. Pattern recognition for road traffic accident severity in Korea. Ergonomics 44(1), 107-117.
- 8、 Miller, T.R., and Blincoe, L.J. (1994), "Incidence and Cost of Alcohol-Involved Crashes in the United States," Accid. Anal. Prev., 26, 583-592.
- 9、 呂奇傑(2001)。演化式類神經網路分類技術於資料探勘上之應用。輔仁大學應用統計學研究所碩士論文，新莊。
- 10、 Lo, A. W., 1986, "Logit Versus Discriminant Analysis- A Specification Test and Application to Corporate Bankruptcies", Journal of Econometrics, 86, 151-178.
- 11、 張紘炬、林顯毓(1995)。台北市長選舉投票傾向的 LOGIT 模式分析。民意研究季刊，192，1-11。
- 12、 張健邦(1995)，多變量分析，三民書局。
- 13、 施孟隆(1999)。Logistic Model 模式應用於信用卡信用風險審核系統之研究-以國內某銀行信用卡中心為例。金融財務月刊，4，85-104。

- 14、 陳家彬、賴怡洵(2001)。台灣地區銀行放款有無擔保之決定因素—Logit模型之實證分析。管理評論，20(1)，129-159。
- 15、 Tu, J. V., 1996 “Advantages and Disadvantages of Using Artificial Neural Networks Versus Logistic Regression for Predicting Medical Outcomes,” Clin. Epidemiol. 49 (11), 1225-1231.
- 16、 鄭瑞英(2003)。老年人在全民健保下中西醫就醫行為研究。中國醫藥大學醫務管理研究所碩士，台中。
- 17、 鄭明智(2004)。以病人用藥安全概念建構醫院處方開立管理制度。長庚大學醫務管理學研究所碩士，台北。
- 18、 陳建勝、王安平(2003)。全民健康保險醫療費用審查制度之研究。朝陽科技大學保險金融管理碩士，台中。
- 19、 劉惠美、雷淑儀(2003)。汽機車單一車輛事故駕駛人死亡勝算模式之研究。國立台北大學統計學系研究所碩士，台北。
- 20、 Kim, K., Lawrence, N., Richardson, J., Li, L., 1995. Personal and behavioral predictors of automobile crash and injury severity. Accident Analysis and Prevention 27 (4), 469 - 481.
- 21、 Mercier, C. R., Shelley, M. C., Rimkus, J., Mercier, J. M., 1997. Age and gender as predictors of injury severity in head-on highway vehicular collisions. In: transportation Research Record 1581, TRB, National Research Council, Washington, DC.
- 22、 張永源(2004)。高雄市大學生機車意外事故流行病學調查。高雄醫學大學公共衛生學研究所碩士，高雄。
- 23、 交通部運輸研究所(2004)，道路交通事故相關資料整合系統雛形建置研究(I)-基本雛形環境之建置。
- 24、 內政部警政署(2003)。台閩地區交通事故原因、傷亡、及車輛損壞-民國九十一年。見：<http://www.npa.gov.tw>
- 25、 張立言(2003)。應用個人違規紀錄預測交通事故發生之研究。國立嘉義大學運輸與物流工程研究所，嘉義。
- 26、 張彩秀(2004)。中部某科技大學學生騎機車行為與事故傷害之相關性研究。弘光學報，43，47-53。
- 27、 Agresti, A., 1996. An Introduction to Categorical Data Analysis, Wiley Series in Probability and Statistics.